10/p86

10/523798
DT12 Rec'd PCT/PTO 09 FEB 2005
PCT/CA2003/001215

TITLE OF THE INVENTION

CONTENT-BASED IMAGE RETRIEVAL METHOD

FIELD OF THE INVENTION

[0001] The present invention relates to digital data retrieval. More specifically, the present invention is concerned with content-based image retrieval.

BACKGROUND OF THE INVENTION

[0002] With advances in the computer technologies and the advent of the World-Wide Web, there has been an explosion in the quantity and complexity of digital data being generated, stored, transmitted, analyzed, and accessed. These data take different forms such as text, sound, images and videos.

[0003] For example, the increasing number of digital images available brings the need to develop systems for efficient image retrieval which can help users locate the needed images in a reasonable time. Some of these retrieval systems use attributes of the images, such as the presence of a particular combination of colors or the depiction of a particular type of event. Such attributes may either be derived from the content of the image or from its surrounding text and data. This leads to various approaches in image retrieval such as content-based techniques and text-based techniques.

[0004] In any case, when an image retrieval system returns the results of a given query, two problems often arise: noise and miss. Noise arises

when images which don't correspond to what the user wants are retrieved by the system. Miss is the set of images corresponding to what the user wants which have not been retrieved. These two problems originate from imperfections at different levels. Indeed, it may not be easy for the user to formulate an adequate query using the available images, either because none of them correspond to what the user wants or because the user lacks sufficient knowledge of imagery details to articulate image features. Also, it has been found difficult to translate the user's needs and specificities in terms of image features and similarity measures.

[0005] More specifically in the case of content-based image retrieval, one can distinguish many ways of formulating queries. Early systems such as QBIC, which is described by **Flicker *et al.*** in "Query by image and video content. The QBIC system" in IEEE Computer Magazine, 28:23--32, 1995, prompt the user to select image features such as color, shape, or texture. Other systems like BLOBWORLD which is described by **Carson *et al.*** in "A system for region-based image indexing and retrieval" from the International Conference on Visual Information Systems, pages 509--516, Amsterdam, 1999, require the user to provide a weighted combination of features.

[0006] However, a drawback of such content-based image retrieval techniques is that it is generally difficult to directly specify the features needed for a particular query, for several reasons. A first of such reasons is that not all users understand the image vocabulary (e.g. contrast, texture, color) needed to formulate a given query. A second reason is that, even if the user is an image specialist, it is not easy to translate the images the user has in mind into a combination of features.

[0007] An alternative approach is to allow the user to specify the features and their corresponding weights implicitly via a visual interface known in the art as "query by example". Via this process, the user can choose images that will participate in the query and weight them according to their resemblance to the images sought. The results of the query can then be refined repeatedly by specifying more relevant images. This process, referred to in the art as “relevance feedback” (RF), is defined **Rui *et al.*** in “Content-based image retrieval with relevance feedback in MARS” from the IEEE International Conference on Image Processing, pages 815-818, Santa Barbara, California, 1997, as the process of automatically adjusting an existing query using information fed back by the user about the relevance of previously retrieved documents.

[0008] Relevance feedback is used to model the user subjectivity in several stages. First, it can be applied to identify the ideal images that are in the user's mind. At each step of the retrieval, the user is asked to select a set of images which will participate in the query; and to assign a degree of relevance to each of them. This information can be used in many ways in order to define an analytical form representing the query intended by the user. The ideal query can then be defined independently from previous queries, as disclosed in "Mindreader: Query databases through multiple examples" in 24th International Conference on Very Large Data Bases, pages 433--438, New York, 1998 by **Ishikawa *et al.*** It can also depend on the previous queries, as in the “query point movement method” where the ideal query point is moved towards positive example and away from negative example. This last method is explained by **Zhang *et al.*** in “Relevance Feedback in Content-Based Image Search” from the 12th International Conference on New Information Technology (NIT) in Beijing, May 2001.

[0009] Relevance feedback allows also to better capture the user's needs by assigning a degree of importance (e.g. weight) to each feature or by transforming the original feature space into a new one that best corresponds to the user's needs and specificities. This is achieved by enhancing the importance of those features that help in retrieving relevant images and reducing the importance of those which do not. Once the importance of each feature is determined, the results are applied to define similarity measures which correspond better to the similarity intended by the user in specific current query.

[0010] The operation of attributing weights to features can also be applied to perform feature selection, which is defined by **Kim *et al.*** in "Feature Selection in Unsupervised Learning via Evolutionary Search" from the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-00), pages 365—369, San Diego, 2000, as the process of choosing a subset of features by eliminating redundant features or those providing little or no predictive information. In fact, after the importance of each feature is determined, feature selection can be performed by retaining only those features which are important enough; the rest being eliminated. By eliminating some features, retrieval performance can be improved because, in a low-dimension feature space, it is easier to define good similarity measures, to perform retrieval in a reasonable time, and to apply effective indexing techniques (for more detail, see "Web Image Search Engines: A Survey. Technical Report N° 276, Université de Sherbrooke, Canada, December 2001, by **Kherfi *et al.***).

[0011] Relevance feedback using positive examples is very well known in the art. For example, **Ishikawa *et al.*** define a quadratic distance function for comparing images. Considering a query consisting of N images, each image represented by an I-dimension feature vector $\vec{x}_n = [x_{n1},...,x_{nI}]^T$,

where T denotes matrix transposition and considering also that the user associates each image participating in the query with a degree of relevance π_n which represents its degree of resemblance with the sought images **Ishikawa *et al.*** compute two parameters, namely the ideal query $\vec{q}=[q_1,...,q_I]^T$ and the ellipsoid distance matrix W, that minimize the quantity D given in Equation (1), which represents the global distance between the query images and the ideal query:

$$D = \sum_{n=1}^{N} \pi_n (\vec{x}_n - \vec{q})^T W (\vec{x}_n - \vec{q}) \qquad (1)$$

A drawback of the method proposed by Ishikawa *et al.* is that it doesn't support the negative example.

[0012] **Rui *et al.*(2)** in "Optimizing Learning in Image Retrieval". IEEE international Conference On Computer Vision and Pattern Recognition, Hilton Head, Sc, USA, 2000 disclose a method where each image is decomposed into a set of *I* features, each of which represented by a vector of reals. $\vec{x}_{ni}$ represents the ith feature vector of the nth query image and π_n the degree of relevance assigned by the user to the nth image. It is assumed also that the query consists of N images. For each feature i, the ideal query vector $\vec{q}_i$, a matrix W_i and scalar weight u_i which minimize the global dispersion of the query images given by Equation (2) are computed. Minimizing the dispersion of the query images aims at enhancing the concentrated features, i.e., features for which example images are close to each other.

$$J = \sum_{i=1}^{I} u_i \sum_{n=1}^{N} \pi_n (\vec{x}_{ni} - \vec{q}_i)^T W_i (\vec{x}_{ni} - \vec{q}_i) \qquad (2)$$

[0013] In "Efficient Indexing, Browsing and Retrieval of Image/Video Content", PhD thesis, Department of Computer Science, University of Illinois at Urbana-Champaign, 1999, **Rui** ***et al*** **(3)** propose to use a similar model but with negative degrees of relevance assigned to negative example images. A drawback of this model, is that it leads to neglect the relevant features of negative example, so that negative example will be confused with positive example.

[0014] It is to be noted that, while many studies have focused on how to learn from user interaction in relevance feedback, few of them evoked the relevance of negative example. However, negative example can be useful for query refinement since it allows to determine the images the user doesn't want in order to discard them. Indeed, **Müller** ***et al.*** shows, in "Strategies for Positive and Negative Relevance Feedback in Image Retrieval.", Technical Report N° 00.01, Computer Vision Group, Computing Center, University of Geneva, 2000, that, using only positive feedback, yields major improvement only at the first feedback step, while improvement is remarkable for the four first steps with positive and negative feedback where the results continuously get better.

[0015] Relevance feedback with negative example may also be useful to reduce noise (undesired images that have been retrieved) and to decrease the miss (desired images that have not been retrieved). Indeed, after the results of a given query are obtained, the user can maintain the positive example images and enrich the query by including some undesired images as negative example. This implies that images similar to those of negative example will be discarded, thus reducing noise. At the same time, the

discarded images will be replaced by others which would have to resemble better what the user wants. Hence, the miss will also be decreased. Furthermore, the user can find, among the recently retrieved images, more images that resemble what the user needs and use them to formulate a new query. Thus, the use of negative example would help to resolve what is called the page zero problem, i.e., that of finding a good query image to initiate retrieval. By mitigating the page zero problem, it has been found that the retrieval time is reduced and the accuracy of the results is improved (see **Kherfi** ***et al.***). It is also to be noted that relevance feedback with negative example is useful when, in response to a user feed-back query, the system returns exactly the same images as in a previous iteration. Assuming that the user has already given the system all the possible positive feedback, the only way to escape from this situation is to choose some images as negative feedback.

[0016] Consider the interpretation of results for content-based image retrieval methods involving negative example, one can distinguish two categories of models. In the first category, the positive example images are selected by the user; however, the negative example images are chosen automatically by the retrieval system among those not selected by the user. In the second category, both positive and negative example images are chosen by the user.

[0017] **Müller** ***et al.*** describe a content-based image retrieval method from the first category. Concerning the initial query, they propose to enrich it by automatically supplying non-selected images as negative example. For refinement, the top 20 images resulting from the previous query as positive feedback are selected. As negative feedback, four of the non-returned images are chosen. The Müller method allows refinement through several feedback

steps; each step aims at moving the ideal query towards the positive example and away from the negative example. More specifically, this is achieved by using the following formula proposed by **Rocchio** in "Relevance Feedback in Information Retrieval" in SMART Retrieval System, Experiments in Automatic Document Processing, pages 323--323, New Jersey, 1971:

$$Q = \frac{\alpha}{n_1}\sum_{i=1}^{n_1} R_i - \frac{\beta}{n_2}\sum_{i=1}^{n_2} S_i \tag{3}$$

where Q is the ideal query, n_1 and n_2 are the numbers of positive and negative images in the query respectively, and R_i and S_i are the features of the positive and negative images respectively. α and β determine the relative weighting of the positive and negative examples. The values $\alpha = 0.65$ and $\beta = 0.35$, which are used for some text-retrieval systems are used (see **Müller *et al.***).

[0018] Since the system selects negative example images automatically, a drawback of systems from the first category, is that using inappropriate images can destroy the query. Indeed, if the system chooses as negative example some images which should rather be considered as positive example, then the relevant features of these images will be discarded, and this will mislead the retrieval process.

[0019] **Vasconcelos *et al.*** in "Learning from User Feedback in Image Retrieval Systems." in Neural Information Processing Systems 12, Denver, Colorado, 1999 disclose a content-based image retrieval methods involving negative example from the second category. More specifically, they propose a Bayesian model for image retrieval, operating on the assumption that the database is constituted of many image classes. When performing retrieval, image classes that assign a high membership probability to positive

example images are supported, and image classes that assign a high membership probability to negative example images are penalized. It is to be noted that the authors consider that the positive and the negative examples have the same relative importance. A drawback of the method and system proposed by Vasconcelos is that it doesn't perform any kind of feature weighting of selection. Indeed, it is well known that the importance of features varies from one user to the other and even from one moment to another for the same user. However, this system considers that all features have the same importance.

[0020] **Picard *et al.*** in "Interactive Learning Using a 'Society of Models' from the IEEE Conference on Computer Vision and Pattern Recognition, pages 447--452, San Francisco, 1996., and in "Modeling user subjectivity in image libraries", Technical Report No. 382, MIT Media Lab Perceptual Computing, 1996, proposed methods involving searching for the set of images similar to positive example, then searching for the set of images similar to negative example; and finally manipulating the two sets in order to obtain the set of images to be returned to the user.

[0021] More specifically, **Picard *et al.*** teach the organization of database images into many hierarchical trees according to individual features such as color and texture. When the user submits a query, comparison using each of the trees are performed; then the resulting sets are combined by choosing the image sets which most efficiently describe positive example, with the condition that these sets don't describe negative example well.

[0022] **Belkin *et al.*** in Rutgers' TREC-6 interactive track experience, from the 6th Text Retrieval Conference, pages 597--610, Gaitherburg, USA, 1998 use a Bayesian probabilistic model in which they assume that the relevant

features of positive example are good, whether or not they are relevant to negative example. Their interpretation of negative example is that the context in which positive example appears is inappropriate to the searcher's problem. They propose to increase the (positive) weight of the relevant features of positive example (irrespective of their appearance in negative example); and to enhance (with negative weights) the relevant features of negative example which don't appear in positive example.

[0023] **Belkin** ***et al.*** consider the negative example at the feature level. They try to identify and enhance the features which help to retrieve images that are at the same time similar to positive example but not similar to negative example. However, enhancing important features of positive example which also appear in negative example can mislead the retrieval process, as will be discussed hereinbelow.

[0024] Finally, **Nastar** ***et al.*** in "Relevance Feedback and Category Search in Image Databases." from the IEEE International Conference on Multimedia Computing and Systems, pages 512--517, Florence, Italy, 1999, and in "Efficient Query Refinement for Image Retrieval." from the IEEE Conference on Computer Vision and Pattern Recognition, pages 547--552, Santa Barbara, 1998, consider an image database made up of relevant images, among which the user chooses positive example, and non-relevant images, among which the user chooses negative example. A probabilistic model is used to estimate the distribution of relevant images and to simultaneously minimize the probability of retrieving non-relevant images. A drawback of such a model is its interpretation of negative example, and how it confuses between negative example images and non-relevant images. In a real database, most images in general are irrelevant to a given query; however, few of them can be used as negative examples without destroying this query..

OBJECTS OF THE INVENTION

[0025] An object of the present invention is therefore to provide improved content-based image retrieval using positive and negative examples.

SUMMARY OF THE INVENTION

[0026] A content-based method for retrieving data files among a set of database files according to the present invention generally aims at defining a retrieval scenario where the user can select positive example images, negative example images, and their respective degrees of relevance. This allows first to reduce the heterogeneity of the dataset on the basis of the positive example, then to refine the results on the basis of the negative example.

[0027] More specifically, in accordance with a first aspect of the present invention, there is provided a content-based method for retrieving data files among a set of database files comprising: providing positive and negative examples of data files; the positive example including at least one relevant feature; providing at least one discriminating feature in at least one of the positive and negative examples allowing to differentiate between the positive and negative examples; for each database file in the set of database files, computing a relevance score based on a similarity of the each database file to the positive example considering the at least one relevant feature; creating a list of relevant files comprising the Nb1 files having the highest similarity score among the set of database files; Nb1 being a predetermined number; for each relevant file in the list of relevant files, computing a discrimination score based on a similarity of the each relevant file to the positive example considering the at least one discriminating feature and on a dissimilarity of the each relevant file to the negative example considering the at least one discriminating feature; and

selecting the Nb2 files having the highest discrimination score among the list of relevant files; Nb2 being a predetermined number.

[0028] In accordance with a second aspect of the present invention, there is provided a content-based method for retrieving images among a set of database images comprising: providing positive and negative example images; the positive example image including at least one relevant feature; providing at least one discriminating feature in at least one of the positive and negative examples allowing to differentiate between the positive and negative example images; for each database image in the set of database images, computing a relevance score based on a similarity of the each database image to the positive example image considering the at least one relevant feature; creating a list of relevant images comprising the Nb1 images having the highest relevance score among the set of database images; Nb1 being a predetermined number; for each relevant image in the list of relevant images, computing a discrimination score based on a similarity of the each relevant image to the positive example image considering the at least one discriminating feature and on a dissimilarity of the each relevant image to the negative example image considering the at least one discriminating feature; and selecting the Nb2 images having the highest discrimination score among the list of relevant images; Nb2 being a predetermined number.

[0029] In accordance with a third aspect of the present invention, there is provided a content-based method for retrieving images among a set of database images, the method comprising: providing positive and negative example images; the positive example image including at least one relevant feature; restricting the set of database images to a subset of images selected among the database images; the images in the subset of images being selected according to their similarity with the positive example based on the at

least one relevant feature; retrieving images in the subset of images according to their similarity with the positive example based on the at least one relevant feature and according to their dissimilarity with the negative example based on at least one discriminating feature between the positive and negative examples; whereby, the images retrieved among the database images corresponding to images similar to the positive example and dissimilar to the negative example.

[0030] A content-based image retrieval method according to the present invention renders unnecessary the computation of the ideal query since it allows to automatically integrate what the user is looking for into similarity measures without the need to identify any ideal point.

[0031] In accordance to a fourth aspect of the present invention, there is provided a content-based system for retrieving images among a set of database images comprising: means for providing positive and negative example images; the positive example image including at least one relevant feature; means for providing at least one discriminating feature in at least one of the positive and negative examples allowing to differentiate between the positive and negative example images; means for computing, for each database image in the set of database images, a relevance score based on a similarity of the each database image to the positive example image considering the at least one relevant feature; means for creating a list of relevant images comprising the Nb_1 images having the highest similarity score among the set of database images; Nb_1 being a predetermined number; means for computing, for each relevant image in the list of relevant images, a discrimination score based on a similarity of the each relevant image to the positive example image considering the at least one discriminating feature and on a dissimilarity of the each relevant image to the negative example image considering the at least one discriminating feature; and means for selecting the

Nb_2 images having the highest discrimination score among the list of relevant images; Nb_2 being a predetermined number.

[0032] In accordance to a fifth aspect of the present invention, there is provided an apparatus for retrieving images among a set of database images, the apparatus comprising; an interface adapted to receive positive and negative example images; the positive example image including at least one relevant feature; a restriction component operable to restrict the set of database images to a subset of images selected among the database images; the images in the subset of images being selected according to their similarity with the positive example based on the at least one relevant feature; a retrieval component operable to retrieve images in the subset of images according to their similarity with the positive example based on the at least one relevant feature and according to their dissimilarity with the negative example based on at least one discriminating feature between the positive and negative examples; whereby, the images retrieved among the database images correspond to images similar to the positive example and dissimilar to the negative example.

[0033] Finally, in accordance to a sixth aspect of the present invention, there is provided a computer readable memory comprising content-based image retrieval logic for retrieving images among a set of database images, the content-based image retrieval logic comprising: image reception logic operable to receive positive and negative example images; the positive example image including at least one relevant feature; restriction logic operable to restrict the set of database images to a subset of images selected among the database images; the images in the subset of images being selected according to their similarity with the positive example based on the at least one relevant feature; and retrieval logic operable to retrieve images in the subset of images according to their similarity with the positive example based on the at least one

relevant feature and according to their dissimilarity with the negative example based on at least one discriminating feature between the positive and negative examples; whereby, the images retrieved among the database images correspond to images similar to the positive example and dissimilar to the negative example.

[0034] Other objects, advantages and features of the present invention will become more apparent upon reading the following non restrictive description of preferred embodiments thereof, given by way of example only with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0035] In the appended drawings:

[0036] Figure 1 is a flowchart illustrating a content-based image retrieval method according to an illustrative embodiment of the present invention;

[0037] Figure 2 is a graph illustrating precision-scope curves for two cases: negative example in two steps according to the method of Figure 1 and negative example in one step according to the prior art;

[0038] Figure 3 is a computer screenshot of a graphical interface displaying sample images related to different subjects and emphasizing different features;

[0039] Figure 4 is a computer screenshot of a query screen from a user-interface allowing a person to characterized example images according to the method of Figure 1;

[0040] Figure 5 is a schematic view illustrating the decomposition of the HIS color space into a set of subspaces and the computation of each subspace's histogram;

[0041] Figure 6 is a graph illustrating a positive average, a negative average, and the resulting overall query average;

[0042] Figure 7 is a graph illustrating the minimization of the global dispersion leading to neglect the relevant features of negative example;

[0043] Figure 8, which is labeled "Prior Art", is a graph illustrating the minimization of the dispersion of positive example, the minimization of negative example and the minimization of the distinction between them according to a method from the prior art;

[0044] Figure 9 is a screenshot illustrating the result following step 106 from the method of Figure 2;

[0045] Figure 10 is a screenshot illustrating the result following step 112 from the method of Figure 2;

[0046] Figure 11 is a graph illustrating precision-scope curves for retrieval with positive example and refinement with negative example; and

[0047] Figure 12 is a table showing the number of iterations needed to locate a given category of images in two cases: using positive example only and using both positive and negative examples according to the method of Figure 2.

DETAILED DESCRIPTION OF THE INVENTION

[0048] A content-based image retrieval method according to the present invention involves relevance feedback using negative examples. The negative examples are considered from the feature point of view, and used to identify the most discriminating features according to a user-given query.

[0049] A content-based image retrieval method according to the present invention makes use of decision rules including characteristic rules and discrimination rules will now be briefly explained. A characteristic rule of a set is an assertion which characterizes a concept satisfied by all or most of the members of this set. For example, the symptoms of a specific disease can be summarized by a characteristic rule. A discrimination rule is an assertion which discriminates a concept of the target set from the rest of the database. For example, to distinguish one disease from others, a discrimination rule should summarize the symptoms that discriminate this disease from others.

[0050] In applying a content-based image retrieval method according to the present invention, it is assumed that positive and negative examples possess some relevant features that are discriminant, i.e., relevant to either positive or negative example or to both but whose values are not the same in positive and in negative examples. In other words, the case in which the relevant features of positive example are the same as those of negative example, with similar values is excluded. Such a case would yield an

ambiguous query. A system implementing a content-based image retrieval method according to the present invention is programmed to reject such a case and to prompt and allow the user to specify new relevant features.

[0051] To implement the above described principle, characteristic rules may first be extracted from positive example images by the identification of their relevant features. More importance should then be given to such features in the retrieval process and images enhancing them should be retrieved. Secondly, discrimination rules can be extracted from the difference between positive example and negative example. Relevant features whose values are not common to positive and negative examples are good discriminators, and hence must be given more importance; conversely, common features are not good discriminators, and must be penalized. However, applying this principle in this manner, may render misleading the retrieval process by neglecting certain relevant features of positive and negative examples, as explained below.

[0052] Before describing in details a content-based image retrieval method according to the present invention, which would solve the problem presented hereinabove, the concept of relevant feature will be define in more detail. A given feature is considered relevant if it helps retrieving the images being sought. This will depend on two factors.

[0053] First, the relevance can be considered with respect to the query. A feature relevant to the query is a feature which is salient in the majority of the query images. A feature to be considered is a feature whose values are concentrated in the query images, and which discriminates well between positive and negative examples, as relevant to the query.

[0054] Second, the relevance of a feature can be considered with respect to the database. If a given feature's values are almost the same for the majority of the database images, then this feature is considered to be not relevant since it doesn't allow to distinguish the sought images from the others; and vice versa. To illustrate this, consider a database in which each image contains an object with a circular shape, but where the color of the object differs from one image to another. In such a database, the shape feature is not interesting for retrieval since it doesn't allow to distinguish between desired and undesired images; however, the color feature is interesting. In other words, a feature in term of which the database is homogeneous is considered not relevant for retrieval; whereas, a feature in term of which the database is heterogeneous is considered relevant.

[0055] In the following, the consequences of neglecting features whose values are common to both positive and negative examples is analyzed. In fact, this depends on the nature of the database. If the database is homogeneous in terms of such features, then neglecting them will not be detrimental since they are not relevant to the database. On the other hand, if the database is heterogeneous in terms of these features, then neglecting them will lead the system to retrieve many undesired images and to miss many desired images.

[0056] From the above, it is clear that common features should be considered to develop a solution that works for any query. However, in some cases, there are not enough common features to be considered alone at a given moment; they must rather be considered together with other features.

[0057] Turning now to Figure 1 of the appended drawings, a content-based image retrieval method 100 according to a first illustrative embodiment of the present invention is illustrated.

[0058] Generally stated the method 100 consists in performing the following steps:

102 - providing a set of database images;

104 - providing positive and negative example images;

106 - for each database image, computing a relevance score based on a similarity of the database image to the positive example image considering relevant features;

108 - creating a list of relevant images comprising the Nb_1 images having the highest relevance score among the set of database images;

110 - providing discriminating features allowing to differentiate between the positive and negative example images;

112 - for each relevant image in the list of relevant images, computing a discrimination score based on the similarity of the relevant image with the positive example image considering the discriminating features and on a dissimilarity of the relevant image with to the negative example image considering the discriminating features; and

114 - selecting the Nb2 images having the highest discrimination score among the list of relevant images.

[0059] It can be useful to described a content-based image retrieval method according to the present invention as including two general steps. In the following, we will refer to the steps of the method 100 using referral

numbers and we will refer to the more general steps using the expressions: first and second general steps.

[0060]. The first general step allows to reduce the heterogeneity of the set of images participating in the retrieval by restricting it to a more homogeneous subset according to positive example relevant features (and thus according to common features also). In this first general step, we enhance all the relevant features of positive example. We rank the database images according to their resemblance to positive example and then retain only the Nb_1 top-ranked images, where Nb_1 is a predetermined number.

[0061] Only images retained in the first general step will participate in the refinement performed in the second general step, where we enhance the discrimination features, i.e., those whose values are not common to positive and negative examples. In this second general step we rank the candidate images according to their similarity to positive example and dissimilarity to negative example, and return to the user only the Nb_2 ($Nb_2 < Nb_1$) top-ranked images. Hence, even if the common features are neglected in the second general step, this will not mislead the retrieval since they were considered in the first general step. As will be presented hereinbelow in more detail, we confirmed experimentally, using a retrieval system implementing the present method, the importance of processing queries with negative example in two steps.

[0062] Figure 2 compares the curves precision-scope for the two techniques: negative example queries processed in two general steps according to a content-based image retrieval according to the present invention versus negative example queries processed in a unique step (in which both positive and negative examples are considered and all images in the database

participate in retrieval) according to methods from the prior art. The ordinate "Precision" represents the average of relevance of retrieved images, and "scope" is the number of retrieved images. It is clear from Figure 1 that when queries containing negative example are considered in one step, the precision of retrieval decreases quickly with the number of retrieved images.

[0063] Before describing each of the steps 102-114 of the method 100, some special cases are important and merit to be mentioned to show that the proposed image retrieval method functions as well. These cases emerge when all the discrimination features come from positive example only or from negative example only. Indeed, if the relevant features of positive example are strictly included in those of negative example and with common values, then applying the proposed principle leads, in the general first step, to enhance the relevant features of positive example (which are the same as the common features) and to retain images looking like it. Then, in the second general step, to enhance the rest of the negative example relevant features and to discard images near to it. On the other hand, if the relevant features of negative example are strictly included in those of positive example and with common values, then applying the proposed principle leads, in the first general step, to enhance the relevant features positive example (which include those of negative example) and to retain images looking like the positive example. Then, in the second general step, to enhance only those features relevant to positive but not to negative example and to re-rank the images according to these features essentially.

[0064] The following will explained how the content base image retrieval method 100 may allow a user to compose a query using negative example only.

[0065] First, we note that, for a given query, the number of non-relevant images is usually much higher than the number of relevant images. In other words, if we know what someone doesn't want, this doesn't inform us sufficiently about what the user wants. For example, if the user gives an image of a car as negative example without giving any positive example, then we cannot know whether the user is looking for images of buildings, animals, persons or other things. Nevertheless, negative example can be used alone in some cases, for instance, to eliminate a subset from a database, for example, when a database contains, in addition to images the user agrees with, other images that the user's culture doesn't tolerate, e.g. nudity images for some persons. In such a case, the user can first eliminate the undesired images by using some of them as negative example; then the user can navigate in, or retrieve from the rest of the database. Concerning the retrieval method, the negative-example-only query will be considered as a positive example query, i.e., the system first searches for images that resemble negative example. Then, when the resulting images (images that the user wants to discard) are retrieved, the system returns to the user the rest of the database rather these images.

[0066] Each of the steps 102-114 of the method 100 will now be described in more detail.

[0067] In step 102, a set of database images is provided to or by a user, among the set of images possibly including images that the user wants to retrieve.

[0068] Then, in step 104, positive and negative example images are provided through interaction between the user and the system implementing the method 100. Of course, the person seeking images having specific

features can alternatively select the example images manually. In that case, the selected images are digitized afterwards.

[0069] The user interaction aims to achieve two main objectives. First, to be able to combine the query images together with their respective degrees of relevance in order to identify what the user is looking for; and to integrate this information in similarity measures. Second, to weight each predetermined feature and its components according to its relevance to the query and the discrimination power it can provide.

[0070] Figure 3 illustrates a graphical interface displaying nine sample images related to different subjects and emphasizing different features. The graphical interface is programmed so as to allow a user to choose additional images from the database before formulating the query. To select an image as an example image (or query image), the user may click on the "Select" button. The system displays a dialog box allowing the user to specify a degree of relevance (see Figure 4). The user-interface illustrated in Figure 4 allows a person to characterize selected example images.

[0071] For each selected images, the possible relevance degrees are

- Very similar: corresponds to the relevance value 2 for a positive example image;
- Similar: corresponds to the relevance value 1 for a positive example image;
- Doesn't matter: the image will not participate in the query;

- Different: corresponds to the relevance value 1 for a negative example image; or
- Very different: corresponds to the relevance value 2 for a negative example image.

[0072] Of course, the relevancy of each image can be characterized with more or less finesse.

[0073] Before explaining in more detail the formulation of relevance feedback, an example of image model and similarity measure will be described. Of course, another image model can alternatively be used.

[0074] To represent images, the hierarchical model proposed by **Rui *et al.*** is used. According to this model, each image, either in the query or in the database, is represented by a set of I features, each of which is a real vector of many components. It has been found that this image model ensures a good modeling of both images and image features, and a reduction in the computation time. According to this hierarchical two-level image model, a distance metric for each level is selected. For feature level, a generalized Euclidean distance function is chosen, as in **Ishikawa *et al.*** If $\vec{x}_{i1}$ and $\vec{x}_{i2}$ are the i^{th} feature vectors of the images x_1 and x_2 respectively, then the distance at this feature level is

$$D_i(\vec{x}_{i1}, \vec{x}_{i2}) = (\vec{x}_{i1} - \vec{x}_{i2})^T W_i(\vec{x}_{i1} - \vec{x}_{i2}) \quad (4)$$

where W_i is a symmetric matrix that allows us to define the generalized ellipsoid distance D_i.

[0075] The choice of this distance metric allows not only to weight each feature's component but also to transform the initial feature space into a space that better models the user's needs and specificities. The global distance between two images x_1 and x_2 is linear and is given by

$$D(x_1, x_2) = \sum_{i=1}^{I} u_i(\vec{x}_{1i} - \vec{x}_{2i})^T W_i(\vec{x}_{1i} - \vec{x}_{2i}) \tag{5}$$

where u_i is the global weight assigned to the i^{th} feature.

[0076] Each image, either in the database or in the query, is represented by a set of 27 feature vectors, computed as follows: First, every pixel in the image is mapped to a point in the three-dimensional (3D) HSI space (Figure 5). This operation consists of computing, for every triple [H,S,I], the number of pixels having the values Hue = H, Saturation = S and Intensity = I. This yields a 3D color histogram that takes up a lot of space and having zeros for most of its values. For example, an image with HSI values ranging between 0 and 255, would yield a histogram containing 256^3 cells, most of which not corresponding to any pixel.

[0077] To reduce the histogram's size, many solutions are possible, such as the spatial repartition of the points of the 3-D histogram, taking into account their respective occurrence frequency, i.e., the number of pixels corresponding to each point in the histogram. However, since the method 100 does not aim at finding the best visual features, a compromise consists in partitioning the space by subdividing the axes H, S and I into three equal intervals each. This gives 3^3 = 27 subspaces, as shown in Figure 5. Each subspace constitutes a feature, and its corresponding vector is computed as

follows. The subspace is subdivided into $2^3 = 8$ sub-subspaces. The sum of the elements of each sub-subspace is computed and the result is stored in the corresponding cell of the feature vector

[0078] Alternatively, the images can be represented using other models.

[0079] In step 106, a relevance score is computed for each database image based on the similarity of the image to the positive example image considering the relevant feature.

[0080] Considering that the user constructs a query composed of N_1 positive example images and their respective relevance degrees π_n^1 for n = 1,...,N_1, as well as N_2 negative example images and their respective relevance degrees π_n^2 for n = 1,...,N_2. (It should be noted that π_n^2 is not the square of π_n; 2 is an index designating the negative example).

[0081] Only the positive examples are considered in step 106. Each relevance feature and its components is enhanced according to its relevance to the positive example. This can be done by introducing the optimal parameters u_i and W_i which minimize $J_{positive}$, the global dispersion of positive example, given in Equation (6).

$$J_{positive} = \sum_{i=1}^{I} u_i \sum_{n=1}^{N_1} \pi_n^1 (\vec{x}_{ni}^1 - \bar{\vec{x}}_i^1)^T W_i (\vec{x}_{ni}^1 - \bar{\vec{x}}_i^1) \qquad (6)$$

where $\bar{\vec{x}}_i^1$ is the weighted average of positive example (see Figure 6), given by

$$\vec{\bar{x}}_i^1 = \frac{\sum_{n=1}^{N_1} \pi_n^1 x_{ni}^1}{\sum_{n=1}^{N_1} \pi_n^1} \tag{7}$$

[0082] An image retrieval method according to the present invention allows to give more weight to features and feature components for which the positive example images are close to each other in the feature space. An informal justification is that if the variance of query images is high along a given axis, any value on this axis is apparently acceptable to the user, and therefore this axis should be given a low weight, and vice versa.

[0083] In step 108, the database images are ranked in increasing order according to a relevance score based on a similarity of each database image to the positive example image considering the relevance features

[0084] More specifically a distance from the positive example average and the Nb_1 top-ranked images is computed are kept for the next steps. This distance is given by Equation (8).

$$D(x_n) = \sum_{i=1}^{I} u_i(\vec{x}_{ni} - \vec{\bar{x}}_i^1)^T W_i(\vec{x}_{ni} - \vec{\bar{x}}_i^1) \tag{8}$$

[0085] If the query contains only negative example images, then the system proceeds initially by a similar procedure, but considering the negative example rather than the positive example. This means that the system computes the ideal parameters which minimize the dispersion of negative example images, ranks the images in increasing order according to their distance from the negative example average, then returns to the user the last-ranked images. If the query contains both positive and negative examples, then the system performs the two steps of retrieval. The parameter computation and

the distance function used in the first step are the same as in the case of a positive-example-only query.

[0086] In the second general step, both positive and negative example images are considered, and the refinement concerns the images retained in the first general step and more specifically in step 108.

[0087] First J_{global}, the global dispersion of the query, including positive and negative example images is defined:

$$J_{global} = \sum_{i=1}^{I} u_i \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \vec{q}_i)^T W_i (\vec{x}_{ni}^k - \vec{q}_i) \tag{9}$$

where k = 1 for positive example and k = 2 for negative example, and where $\vec{q}_i$, given in Equation (10), is the weighted average of all query images for the ith feature (see Figure 7).

$$\vec{q}_i = \frac{\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k \vec{x}_{ni}^k}{\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k} \tag{10}$$

[0088] In **Rui *et al.* (2)**, it is proposed to allocate negative degrees of relevance to negative example images and to compute the parameters which minimize the same expression of Equation (9). The consequences of such an approach, which is not adopted in a content-based image retrieval method according to the present invention, will now be considered in order to emphasis the differences such an approach and the one used in the method 100. If positive example are considered separately from negative example in Equation (9), then:

$$J_{global} = \sum_{i=1}^{I} u_i \sum_{n=1}^{N_1} \pi_n^1 (\vec{x}_{ni}^1 - \vec{q}_i)^T W_i (\vec{x}_{ni}^1 - \vec{q}_i) + \sum_{i=1}^{I} u_i \sum_{n=1}^{N_2} \pi_n^2 (\vec{x}_{ni}^2 - \vec{q}_i)^T W_i (\vec{x}_{ni}^2 - \vec{q}_i) \quad (11)$$

[0089] **Rui *et al.* (2)** choose $\pi_n^1 > 0$ for n = 1,...,N_1 and $\pi_n^2 < 0$ for n = 1,...,N_2, yielding:

$$J_{global} = \sum_{i=1}^{I} u_i \sum_{n=1}^{N_1} \pi_n^1 (\vec{x}_{ni}^1 - \vec{q}_i)^T W_i (\vec{x}_{ni}^1 - \vec{q}_i) - \sum_{i=1}^{I} u_i \sum_{n=1}^{N_2} |\pi_n^2| (\vec{x}_{ni}^2 - \vec{q}_i)^T W_i (\vec{x}_{ni}^2 - \vec{q}_i) \quad (12)$$

[0090] where $|\pi_n^2|$ designates the absolute value of π_n^2. Equation (12) shows that the global dispersion J_{global} is the dispersion of positive example minus the dispersion of negative example. Hence, by minimizing the global dispersion, even if **Rui *et al.* (2)** move the global query average q (with which they compare their images) towards positive example and away from negative example, two problems emerge.

[0091] First, minimizing the global dispersion will lead to minimize the dispersion of positive example, but with respect to the global query average q rather than the positive example average $\bar{x}_1$. This will not give an optimal minimization of the positive example dispersion; and hence, the relevant features of positive example will not be given enough importance.

[0092] Second, minimizing the global dispersion will lead to maximize the dispersion of negative example. This implies that they neglect the relevant features of negative example. Hence, their retrieval system will not be able to discard the undesired images. This is illustrated in Figure 8.

[0093] The weights u_i and W_i are introduced to give more importance to the relevant features of either positive or negative example which allow to distinguish well between them. In other words, via u_i and W_i, weights are attributed to features and the feature space is transformed into a new space in which positive example images are as close as possible, negative example images are as close as possible, and positive example is as far as possible from negative example (see Figure 7). These objectives are translated into a mathematical formulation, by first distinguishing positive example images from negative example images in the global dispersion formula of Equation (9). For each feature i, the weighted average of positive example images $\bar{\vec{x}}_i^1$ is recalled and the weighted average of negative example images $\bar{\vec{x}}_i^2$ in Equations (13) and (14) respectively is defined.

$$\bar{\vec{x}}_i^1 = \frac{\sum_{n=1}^{N_1} \pi_n^1 x_{ni}^1}{\sum_{n=1}^{N_1} \pi_n^1} \tag{13}$$

$$\bar{\vec{x}}_i^2 = \frac{\sum_{n=1}^{N_2} \pi_n^2 x_{ni}^2}{\sum_{n=1}^{N_2} \pi_n^2} \tag{14}$$

[0094] By introducing $\bar{\vec{x}}_i^1$ and $\bar{\vec{x}}_i^2$ into Equation (9), one can rewrite it as follows:

$$J_{global} = \sum_{i=1}^{I} u_i \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k \left[(\vec{x}_{ni}^k - \bar{\vec{x}}_i^k) + (\bar{\vec{x}}_i^k - \vec{q}_i)\right]^T W_i \left[(\vec{x}_{ni}^k - \bar{\vec{x}}_i^k) + (\bar{\vec{x}}_i^k - \vec{q}_i)\right] \tag{15}$$

[0095] Developing Equation (15) gives

$$J_{global} = \sum_{i=1}^{I} u_i \Big[\Big(\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k)^T W_i (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k) \Big) + \Big(\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k)^T W_i (\vec{\bar{x}}_i^k - \vec{q}_i) \Big)$$
$$+ \Big(\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{\bar{x}}_i^k - \vec{q}_i)^T W_i (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k) \Big) + \Big(\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{\bar{x}}_i^k - \vec{q}_i)^T W_i (\vec{\bar{x}}_i^k - \vec{q}_i) \Big) \Big] \quad (16)$$

[0096] It can easily be shown that the second and third parts of Equation (16) are zero. For example, the second part

$$\textstyle\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k)^T W_i (\vec{\bar{x}}_i^k - \vec{q}_i) = \sum_{k=1}^{2} \Big[\Big(\sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k)^T \Big) W_i (\vec{\bar{x}}_i^k - \vec{q}_i) \Big]$$

$$= \textstyle\sum_{k=1}^{2} \Big[\Big((\sum_{n=1}^{N_k} \pi_n^k \vec{x}_{ni}^k) - (\sum_{n=1}^{N_k} \pi_n^k) \vec{\bar{x}}_i^k \Big)^T W_i (\vec{\bar{x}}_i^k - \vec{q}_i) \Big] = 0$$

since, according to Equations (13) and (14),

$$\textstyle\sum_{n=1}^{N_k} \pi_n^k x_{ni}^k - (\sum_{n=1}^{N_k} \pi_n^k) \vec{\bar{x}}_i^k = 0$$

[0097] Thus, Equation (17) can be written as follows:

$$J_{global} = \Big[\sum_{i=1}^{I} u_i \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k)^T W_i (\vec{x}_{ni}^k - \vec{\bar{x}}_i^k) \Big] + \Big[\sum_{i=1}^{I} u_i \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{\bar{x}}_i^k - \vec{q}_i)^T W_i (\vec{\bar{x}}_i^k - \vec{q}_i) \Big] = A + R \quad (17)$$

[0098] The first term "A" expresses the positive example internal dispersion, i.e., how close positive example images are to each other, added to the negative example internal dispersion, i.e., how close negative example images are to each other. The second term "R" expresses the distance between the two sets, i.e., how far positive example is from negative example.

[0099] By distinguishing the intra dispersion "A" from the inter dispersion "R", it is now clearer how one can formulate the above-identified objectives in a mathematical problem. In fact, one want to compute the model

parameters, namely u_i and W_i, which minimize the intra dispersion "A" and maximize the inter dispersion "R". Several combinations of A and R are possible.

[00100] The parameters which minimize the ratio $\frac{A}{R}$, assuming that $R \neq 0$ will be computed. In the case of $R = 0$, the positive example and the negative example are not distinguishable and the query is ambiguous. In such case, the query is rejected and the user is asked to formulate a new one. Furthermore, to avoid numerical stability problems, the following two constraints are introduced: $\sum_{i=1}^{I} \frac{1}{u_i} = 1$ and $\det(W_i)=1$ for all i=1,...,I. By using Lagrange multipliers, the optimal parameters u_i and W_i must minimize the quantity L given in Equation (18).

$$L = \frac{A}{R} - \lambda\Big(\sum_{i=1}^{I} \frac{1}{u_i} - 1\Big) - \sum_{i=1}^{I} \lambda_i\big(det(W_i) - 1\big) \tag{18}$$

where

$$A = \sum_{i=1}^{I} u_i \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)^T W_i (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k) \tag{19}$$

and

$$R = \sum_{i=1}^{I} u_i \sum_{k=1}^{2} \tilde{\pi}^k (\bar{\vec{x}}_i^k - \vec{q}_i)^T W_i (\bar{\vec{x}}_i^k - \vec{q}_i) \tag{20}$$

$\tilde{\pi}^1$ denotes the sum of positive example relevance degrees, i.e., $\tilde{\pi}^1 = \sum_{n=1}^{N_1} \pi_n^1$ and $\tilde{\pi}^2$ denotes the sum of negative example relevance degrees, i.e., $\tilde{\pi}^2 = \sum_{n=1}^{N_2} \pi_n^2$.

[00101] The optimization problem in order to obtain the optimal parameters u_i and W_i will now be resolved.

[00102] It is to be noted first that the relative importance of positive and negative examples are to be determined, i.e., $\tilde{\pi}^1$ with respect to $\tilde{\pi}^2$. Some image retrieval systems, such as the one described by **Müller *et al.*** adopt the values used by certain text retrieval systems which are 0.65 for positive example and 0.35 for negative example. Other systems such as the one described by **Vasconcelos *et al.*** assume that positive example and negative example have the same importance. In the method 100, the latter choice is adopted because it allows some simplifications in the derivation of the problem. Furthermore, all the user-given relevance degrees are normalized so that $\tilde{\pi}^1 + \tilde{\pi}^2 = 1$.

[00103] To obtain the optimal solution for W_i, the partial derivative of L with respect to $w_{i_{rs}}$ for r,s=1,…,H_i, is taken where H_i is the dimension of the i^{th} feature and $w_{i_{rs}}$ is the rs^{th} element of W_i, i.e., $W_i = [w_{i_{rs}}]$, yielding

$$\frac{\partial L}{\partial w_{i_{rs}}} = \frac{R\frac{\partial A}{\partial w_{i_{rs}}} - A\frac{\partial R}{\partial w_{i_{rs}}}}{R^2} - \lambda_i \frac{\partial det(W_i)}{\partial w_{i_{rs}}} \tag{21}$$

where

$$\frac{\partial A}{\partial w_{i_{rs}}} = u_i \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (x_{ni_r}^k - \bar{x}_{i_r}^k)(x_{ni_s}^k - \bar{x}_{i_s}^k) \tag{22}$$

and

$$\frac{\partial R}{\partial w_{i_{rs}}} = u_i \sum_{k=1}^{2} \tilde{\pi}^k (\bar{x}_{i_r}^k - q_{i_r})(\bar{x}_{i_s}^k - q_{i_s}) \tag{23}$$

[00104] Before computing $\frac{\partial L}{\partial w_{i_{rs}}}$, it is to be noted that $\det(W_i) = \sum_{r=1}^{H_i} (-1)^{r+s} w_{i_{rs}} \det(W_{i_{rs}})$, where $\det(W_{i_{rs}})$ is the rs^{th} minor of W_i obtained by eliminating the r^{th} row and the s^{th} column of $\det(W_i)$. Hence,

$$\frac{\partial det(W_i)}{\partial w_{i_{rs}}} = (-1)^{r+s} det(W_{i_{rs}}) \tag{24}$$

By substituting Equations (19), (20) and (21) in (18), we obtain

$$\frac{\partial L}{\partial w_{i_{rs}}} = 0 \Leftrightarrow$$

$$R\left[u_i \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (x_{ni_r}^k - \bar{x}_{i_r}^k)(x_{ni_s}^k - \bar{x}_{i_s}^k)\right] - A\left[u_i \sum_{k=1}^{2} \bar{\pi}^k (\bar{x}_{i_r}^k - q_{i_r})(\bar{x}_{i_s}^k - q_{i_s})\right] - R^2 \lambda_i (-1)^{r+s} det(W_{i_{rs}}) = 0 \Leftrightarrow$$

$$det(W_{i_{rs}}) = \frac{u_i}{(-1)^{r+s} \lambda_i R^2} \left[R \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (x_{ni_r}^k - \bar{x}_{i_r}^k)(x_{ni_s}^k - \bar{x}_{i_s}^k) - A \sum_{k=1}^{2} \bar{\pi}^k (\bar{x}_{i_r}^k - q_{i_r})(\bar{x}_{i_s}^k - q_{i_s})\right] \tag{25}$$

[00105] Now consider the matrix $W_i^{-1} = [w_{i_{rs}}^{-1}]$, the inverse matrix of W_i (provided that W_i is invertible). To obtain the value of each component $w_{i_{rs}}^{-1}$, the determinant method for matrix inversion is used to obtain

$$w_{i_{rs}}^{-1} = \frac{(-1)^{r+s} det(W_{i_{rs}})}{det(W_i)}$$

Knowing that $\det(W_i)=1$ yields

$$w_{i_{rs}}^{-1} = (-1)^{r+s} det(W_{i_{rs}}) \tag{26}$$

[00106] In Equation (26), $\det(W_{i_{rs}})$ is replaced by its value from Equation (25) to obtain

$$w_{i_{rs}}^{-1} = \frac{1}{\gamma}\left[R\sum_{k=1}^{2}\sum_{n=1}^{N_k}\pi_n^k(x_{ni_r}^k - \bar{x}_{i_r}^k)(x_{ni_s}^k - \bar{x}_{i_s}^k) - A\sum_{k=1}^{2}\bar{\pi}^k(\bar{x}_{i_r}^k - q_{i_r})(\bar{x}_{i_s}^k - q_{i_s})\right] \tag{27}$$

where

$$\gamma = \frac{\lambda_i R^2}{u_i}$$

[00107] Equation (27) can also be written in matrix form as

$$W_i^{-1} = \frac{1}{\gamma}C_i \tag{28}$$

where C_i is the matrix $[c_{i_{rs}}]$ such that

$$c_{i_{rs}} = R\sum_{k=1}^{2}\sum_{n=1}^{N_k}\pi_n^k(x_{ni_r}^k - \bar{x}_{i_r}^k)(x_{ni_s}^k - \bar{x}_{i_s}^k) - A\sum_{k=1}^{2}\bar{\pi}^k(\bar{x}_{i_r}^k - q_{i_r})(\bar{x}_{i_s}^k - q_{i_s}) \tag{29}$$

[00108] The value of γ will now be computed independently from λ which is an unknown parameter. Equation (28) can be written as follows:

$$W_i^{-1} = \frac{1}{\gamma}C_i \Leftrightarrow C_i = \gamma W_i^{-1} \Rightarrow det(C_i) = \gamma^{H_i} det(W_i^{-1})$$

but since $\det(W_i^{-1}) = 1$, then $\gamma = (\det(C_i))^{\frac{1}{H_i}} C_i^{-1}$. Finally, the optimal solution for W_i is given by Equation (30)

$$W_i = \gamma C_i^{-1} = (det(C_i))^{\frac{1}{H_i}} C_i^{-1} \tag{30}$$

where the components of C_i are given by Equation (29).

[00109] In the following, the effect of the dispersion of positive and negative examples on the components of W_i will be considered. First, Equation (29) can be rewritten in a matrix form, as follows:

$$C_i = RCova_i - ACovr_i \tag{31}$$

where $Cova_i$ is the sum of intra covariance matrices for the i[th] feature, i.e., $Cova_i = [\text{cov}\, a_{i_{rs}}]$ such that

$$cova_{i_{rs}} = \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (x_{ni_r}^k - \bar{x}_{i_r}^k)(x_{ni_s}^k - \bar{x}_{i_s}^k)$$

and $Covr_i$ is the inter covariance matrix for the i[th] feature, i.e., $Covr_i = [\text{cov}\, r_{i_{rs}}]$ such that

$$covr_{i_{rs}} = \sum_{k=1}^{2} \tilde{\pi}^k (\bar{x}_{i_r}^k - q_{i_r})(\bar{x}_{i_s}^k - q_{i_s})$$

[00110] Now, considering Equation (31), where the values of "A" and "R" are set since they concern all the features. If the intra dispersion is high relative to the inter dispersion, and hence the elements of $Cova_i$ are important relative to the elements of $Covr_i$ then, according to Equation (31), the values of the components of C_i will be important. But since $W_i = \gamma C_i^{-1}$ (Equation(30)), it follows that the values of $w_{i_{rs}}$ will be small; and consequently, the i[th] feature's components will be given low weights. On the other hand, if the intra dispersion

is low relative to the inter dispersion for the i^{th} feature, by a similar line of reasoning, one can see that this feature's components will be given high weights. This behavior of W_i fulfills the objective of enhancing discriminant features against other ones.

[0.0111] Taking the partial derivative of L with respect to u_i allows to obtain the optimal solution for u_i.

$$\frac{\partial L}{\partial u_i} = \frac{R\frac{\partial A}{\partial u_i} - A\frac{\partial R}{\partial u_i}}{R^2} + \frac{\lambda}{u_i^2} \quad (32)$$

where

$$\frac{\partial A}{\partial u_i} = \sum_{k=1}^{2}\sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)^T W_i (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k) \quad (33)$$

and

$$\frac{\partial R}{\partial u_i} = \sum_{k=1}^{2} \tilde{\pi}^k (\bar{\vec{x}}_i^k - \vec{q}_i)^T W_i (\bar{\vec{x}}_i^k - \vec{q}_i) \quad (34)$$

[00112] By substituting Equations (33) and (34) in (32), we obtain

$$\frac{\partial L}{\partial u_i} = 0 \Leftrightarrow R\Big[\sum_{k=1}^{2}\sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)^T W_i (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)\Big] - A\Big[\sum_{k=1}^{2} \tilde{\pi}^k (\bar{\vec{x}}_i^k - \vec{q}_i)^T W_i (\bar{\vec{x}}_i^k - \vec{q}_i)\Big] + \frac{\lambda R^2}{u_i^2} = 0 \quad (35)$$

[00113] Both sides of Equation (35) are multiplied by u_i, to obtain:

$$u_i f_i + \frac{\lambda R^2}{u_i} = 0 \quad (36)$$

where

$$f_i = R\Big[\sum_{k=1}^{2}\sum_{n=1}^{N_k} \pi_n^k(\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)^T W_i(\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)\Big] - A\Big[\sum_{k=1}^{2} \tilde{\pi}^k(\bar{\vec{x}}_i^k - \vec{q}_i)^T W_i(\bar{\vec{x}}_i^k - \vec{q}_i)\Big] \quad (37)$$

[00114] Now, to get rid of the unknown parameter λ, a relation, independent of λ, between u_i and any u_j is sought. First λ can be computed directly from Equation (36) as follows:

$$\lambda = -\frac{f_i u_i^2}{R^2}\ \forall i \quad (38)$$

[00115] Second, taking the sum on i of Equation (36) gives $\sum_{j=1}^{I} u_j f_j + \lambda R^2 \sum_{j=1}^{I}\frac{1}{u_j} = 0$, but since $\sum_{i=1}^{I}\frac{1}{u_i} = 1$, then $\sum_{j=1}^{I} u_j f_j + \lambda R^2 = 0$. It follows that

$$\lambda = \frac{-\sum_{i=j}^{I} u_j f_j}{R^2} \quad (39)$$

[00116] Equations (32) and (33) imply that for every feature i

$$f_i u_i^2 = \sum_{j=1}^{I} u_j f_j \quad (40)$$

[00117] It follows from Equation (40) that $f_1 u_1^2 = f_2 u_2^2 = \ldots = f_i u_i^2 = f_I u_I^2$.

[00118] Hence,

$$u_j = u_i \sqrt{\frac{f_i}{f_j}} \quad \forall j \tag{41}$$

[00119] Finally, to obtain the optimal solution of u_i, u_j is replaced in Equation (40) by its value from Equation (41), yielding:

$$f_i u_i^2 = \sum_{j=1}^{I} \left(u_i \sqrt{\frac{f_i}{f_j}} f_j \right) \Leftrightarrow f_i u_i = \sum_{j=1}^{I} \sqrt{f_i f_j}$$

$$\Leftrightarrow u_i = \frac{\sum_{j=1}^{I} \sqrt{f_j}}{\sqrt{f_i}} \tag{42}$$

[00120] The optimal solution for u_i is given by Equation (42), where f_i is defined by Equation (37).

[00121] The influence of the dispersion of positive and negative examples on the value of each u_i will now be considered First, f_i can be written in Equation (37) as

$$f_i = RFa_i - AFr_i \tag{43}$$

where

$$Fa_i = \sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)^T W_i (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k) \tag{44}$$

and

$$Fr_i = \sum_{k=1}^{2} \tilde{\pi}^k (\bar{\vec{x}}_i^k - \vec{q}_i)^T W_i (\bar{\vec{x}}_i^k - \vec{q}_i) \tag{45}$$

[00122] It is assumed that A and R have constant values since they depend on all the features. If, for the i^{th} feature, the intra dispersion is high relative to the inter dispersion, then the quantity Fa_i will gain in importance relative to the quantity Fr_i. According to Equation (43), this will increase the value of f_i. Moreover, Equation (42) shows that when f_i increases, u_i decreases; and hence, the i^{th} feature will be given a low weight. Conversely, if, for the i^{th} feature, the intra dispersion is low relative to the inter dispersion, then, by a similar line of reasoning, we find that the i^{th} feature will be given a high weight. Therefore, the optimal value that is found for u_i fulfills the objective of enhancing the relevant discriminant features against others.

[00123] In brief, the input to step 112 consists of positive example images, negative example images and their respective relevance degrees. A partial result of step 112 includes the optimal parameters W_i and u_i. These parameters are computed according to Equations (30) and (42), respectively. The computation of these parameters requires the computation of $\bar{\bar{x}}_i^1$, $\bar{\bar{x}}_i^2$, $\vec{q}_i$, f_i, A and R according to Equations (13), (14), (10), (37), (19) and (20), respectively. The algorithm is iterative since the computation of W_i and u_i depends on A and R, and the computation of A and R depends on W_i and u_i. The fixed point method is used to perform the computation of W_i and u_i. An initialization step is required, in which we adopt the following values:

[00124] - W_i is initialized with the diagonal matrix

$$\begin{pmatrix} \frac{1}{\sigma_{i_1}} & \cdot\ \cdot & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot\ \cdot & \cdot \\ 0 & \cdot\ \cdot & \frac{1}{\sigma_{i_{H_i}}} \end{pmatrix}$$

where

$$\sigma_{ir} = \sqrt{\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k \left(x_{ni_r}^k - q_{i_r}\right)^2}$$

is the standard deviation of the r^{th} component of the i^{th} feature computed for the full set of query images.

[00125] - The parameter u_i is initialized with a kind of dispersion given by

$$u_i = \frac{\sum_{j=1}^{I} \sqrt{f_j}}{\sqrt{f_i}}$$

where

$$f_i = \frac{\sum_{k=1}^{2} \sum_{n=1}^{N_k} \pi_n^k (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)^T W_i (\vec{x}_{ni}^k - \bar{\vec{x}}_i^k)}{\sum_{k=1}^{2} \bar{\pi}^k (\bar{\vec{x}}_i^k - \vec{q}_i)^T W_i (\bar{\vec{x}}_i^k - \vec{q}_i)}$$

[00126] The computation of W_i requires the inversion of the matrix C_i. However, in the case of $(N_1+N_2)<H_i$, C_i is not invertible. **Ishikawa *et al.*** suggest proceeding by singular value decomposition (SVD) to obtain the pseudo inverse matrix. However, this solution doesn't give a satisfactory result,

especially when $(N_1+N_2)$is far less than H_i as pointed out by **Rui *et al.***, who propose, in the case of a singular matrix, to replace W_i by a diagonal matrix whose elements are the inverse of the standard deviation, i.e., $w_{i_{rs}} = \frac{1}{\sigma_{i_s}}$ if r = s and $w_{i_{rs}} = 0$ elsewhere.

[00127] In step 112, W_i is replaced by a diagonal matrix whose elements are the inverse of the diagonal elements of the matrix C_i, i.e.,

$$W_i = \begin{pmatrix} w_{i_{11}} & \cdot & \cdot & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot \\ 0 & \cdot & \cdot & w_{i_{H_iH_i}} \end{pmatrix}$$

where $w_{i_{ss}} = \frac{1}{c_{i_{ss}}}$ and $c_{i_{ss}}$ can be obtained by setting r = s in Equation (26).

[00128] In step 114, the relevant images obtained in step 108 are ranked according to a discriminating score based on their closeness to the positive example and their farness from the negative example. The comparison function is given by Equation (44). Finally, the system returns the Nb_2 top-ranked images to the user.

$$D(x_n) = \sum_{i=1}^{I} u_i(\vec{x}_{ni} - \vec{\bar{x}}_i^1)^T W_i(\vec{x}_{ni} - \vec{\bar{x}}_i^1) - \sum_{i=1}^{I} u_i(\vec{x}_{ni} - \vec{\bar{x}}_i^2)^T W_i(\vec{x}_{ni} - \vec{\bar{x}}_i^2) \quad (46)$$

Experimental results and performance evaluation

[00129] Tests were performed on 10 000 images from The Pennsylvania State University images database, which is described by J. Li, J.Z. Wang and G. Wiederhold in both "IRM: Integrated region matching for image retrieval." From the 2000 ACM Multimedia Conference, pages 147--156, San Jose, USA, 2000. and "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries." from IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9):947--963, 2001. This database contains images related to different subjects, emphasizing different features, and taken under different illumination conditions. For each image, the set of features is computed as explained above. Many tests were performed for retrieval and refinement. Even when positive and negative examples are not readily distinguishable, the method according to the present invention succeeded in identifying discrimination features and sorting the resulting images according to these features.

[00130] Figure 9 shows an example of retrieval with positive example only. Figure 10 shows and example of retrieval with positive and negative examples.

[00131] In the first example, two images participated in the query as positive example. Both of these images contain a green tree under the blue sky (5095.ppm and 5118.ppm). Figure 9 shows the top nine returned images. It is to be noted that the two query images are returned in the top positions. There are also some other images containing trees under the sky, but including noise consisting of three images of a brown bird on a green tree under the blue sky (5523.ppm, 5522.ppm, 5521.ppm). At the same time, there have been miss,

because the database contains other images (not shown) of trees under the sky that have not been retrieved.

[00132] According to the second example, a refinement has been applied to the results of the first example. Hence, we use the same images (5095.ppm and 5118.ppm) as positive example, while an image of a bird on a tree under the sky is chosen as negative example (image 5521.ppm of Figure 8). Figure 9 shows that images of birds are discarded (the noise reduced) and that more images of trees under the sky are retrieved (the miss decreased).

Performance evaluation

[00133] In order to validate the proposed relevance feedback technique, a performance evaluation of a retrieval system implementing a method according to the present invention has been has been performed. The evaluation was based on comparison between the use of positive example only and the use of both positive and negative examples. To perform any evaluation in the context of image retrieval, two main issues emerge: the acquisition of ground truth and the definition of performance criteria. For ground truth, human subjects were used: three persons participated in all the experiences described hereinbelow. The performance criteria, Precision Pr and Recall Re, described by John R. **Smith** in "Image Retrieval Evaluation." From the IEEE Workshop on Content-based Access of Image and Video Libraries, 1998 were used.

[00134] In their simplest definition, Precision is the proportion of retrieved images that are relevant, i.e., number of retrieved images that are relevant on the number of all retrieved images; and Recall is the proportion of relevant images that are retrieved, i.e., number of relevant images that are retrieved on the number of all relevant images in the database. **Smith** drew up the precision-recall curve Pr=f(Re); however, it has been observed that this

measure is less meaningful in the context of image retrieval since Recall is consistently low. Furthermore, it is believed that it is often difficult to compute Recall, especially when the size of the image database is big; because this requires to know, for each query, the number of relevant images in a the whole database. Another problem with Recall, is that it depends strongly on the choice of the number of images to return to the user. If the number of relevant images in the database is bigger than the number of images returned to the user, then the recall will be penalized. A more expressive curve which is the precision-scope curve Pr=f(Sc), as described by **Huang *et al.***, "Image Indexing using Color Correlogram." From the IEEE Conference on Computer Vision and Pattern Recognition, 1997, has been used. Scope Sc is the number of images returned to the user, and hence the curve Pr=f(Sc) depicts the precision for different values of the number of images returned to the user. Since these performance criteria are believed to be well known in the art, they will not be described herein in further detail.

[00135] Two experiences were carried out, each of which aiming to measure a given aspect of our model. The first experience aims to measure the improvement, with negative example, in the relevance of retrieved images. The second experience aims to measure the improvement, with negative example, in the number of iterations needed to locate a given category of images.

First experience

[00136] As mentioned above, the goal of the first experience is to measure the contribution of negative example in the improvement of the relevance of retrieved images. Each human subject participating in the experience was asked to formulate a query using only positive example and to give a goodness score to each retrieved image, then to refine the results using negative example and to give a goodness score to each retrieved image. The

possible scores are 2 if the image is good, 1 if the image is acceptable, and 0 if the image is bad. Each subject repeated the experience five times by specifying a new query each time. Precision was computed as follows: Pr = the sum of degrees of relevance for retrieved images / the number of retrieved images. Figure 11 illustrates a comparison between the curves Pr=f(Sc) in the two cases: retrieval with positive example and refinement with negative example.

[00137] The experiences shows that, in average, when negative example is introduced, the improvement in precision is about 20 %. In fact, the improvement varies from one query to another, because it depends on other factors such as the choice of a meaningful negative example and the constitution of the database. If, for a given query, the database contains a little number of relevant images, most of which have been retrieved in the first step, then the introduction of negative example or any other technique will not be able to bring any notable improvement.

Second experience

[00138] The second experience aims at measuring the improvement in the number of refinement iterations needed to locate a given category of images, as well as the role of negative example in resolving the page zero problem (finding a good image to initiate the retrieval). Each of our human subjects was shown a set of images that are relatively similar to each other with respect to the color. None of the showed images appear in the set of images the subjects can use to formulate the initial query. Each subject is asked to locate at least one of the showed images using only positive example, and to count the number of iterations; then to restart the experience but using both positive and negative examples, and to count the number of iterations. This experience was repeated four times and the results are given in Figure 12. S1,

S2 and S3 designate respectively the three human subjects who participated in the experiments. PE means positive example and NE means negative example. Each entry in the table gives the number of iterations needed to locate the searched images.

[00139] It has been found that when they used both positive and negative examples, the subjects succeeded in all the experiences; however, when they used only positive example, some of them failed in certain experiences to locate any sought image. In Experience 2.2 and Experience 2.4, at least one subject was unable to locate any sought image using positive example only. This is because, in a given iteration, all the retrieved images fall into an undesired category, and the formulation of the next-iteration query using any of these images leads to retrieve images belonging to the same category. The user can loop indefinitely, but will not be able to escape this situation by using positive example only. The second observation is that the use of negative example reduces appreciably the number of iterations. If one computes the average number of iterations among the successful experiences (2.1 and 2.3), one finds 5.83 when only positive example is used, and 2.33 when both positive and negative examples are used. This experience shows clearly the role of negative example in mitigating the page zero problem. Indeed, after having obtaining at least one of the sought images, the user can use it to formulate a new query, and hence to retrieve more sought images.

[00140] A content-based image retrieval method according to the present invention allows to take into account the user's needs and specificities, which can be identified via relevance feedback. It has been shown that the use of positive example only isn't always sufficient to determine what the user is looking for. This can be seen especially when all the candidate images to participate in the query appear in an inappropriate context or contain, in

addition to the features the user is looking for, features or objects that the user doesn't want to retrieve.

[00141] It is to be noted that the present model is not limited to image retrieval but can be adapted and applied to any retrieval process with relevance feedback. For example, a method according to the present invention can be used any process of retrieval such as retrieval of text, sound, and multimedia.

[00142] Although the present invention has been described hereinabove by way of preferred embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention.